

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE September 1995		3. REPORT TYPE AND DATES COVERED Professional Paper	
4. TITLE AND SUBTITLE A CHARACTERIZATION OF PROBABILISTIC ENTAILMENT IN ADAMS' LOGIC OF CONDITIONALS				5. FUNDING NUMBERS PR: ZW84 PE: 0601152N WU: DN304020	
6. AUTHOR(S) D. Bamber					
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Command, Control and Ocean Surveillance Center (NCCOSC) RDT&E Division San Diego, CA 92152-5001				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Office of Naval Research Independent Research Program (IR) 800 North Quincy Street Arlington, VA 22217-5660				10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES					
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.				12b. DISTRIBUTION CODE DTIC SELECTED NOV 27 1995 F	
13. ABSTRACT (Maximum 200 words) Ernest Adams' logic of conditionals is useful in expert systems because it can be used to reason about imperfect generalizations, i.e., generalizations of the form: <i>In nearly all instances where α is true, β is true.</i> This paper defines a notion of importance and then shows that Adams' logic is equivalent to a logic for reasoning about generalizations of the form: <i>In all important instances where α is true, β is true.</i> A new method of proving results in Adam's logic—a method based upon the importance notion—is described. This method of proof is often quite easy, particularly since Venn diagrams can be used to help find proofs. An example using this new method of proof is presented. The example shows that Adams' logic justifies a common method of conflict resolution in expert systems, namely, the specificity-wins method. Finally, it is shown that Adams' logic can be used to reason about a topic seemingly unrelated to imperfect generalizations: the choices made by rational agents.					
14. SUBJECT TERMS Adams' Logic Notion of Importance					
15. NUMBER OF PAGES				16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED		18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED		19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	
				20. LIMITATION OF ABSTRACT SAME AS REPORT	

19951122 025

~~Published in Proceedings of the 1994 Symposium on Command and Control Research and Decision Aids, pp. 637-642, July 1994.~~

A Characterization of Probabilistic Entailment in Adams' Logic of Conditionals

Donald Bamber

NCCOSC RDTE DIV 44209
53355 RYNE ROAD RM 222
SAN DIEGO CA 92152-7252

Abstract

Ernest Adams' logic of conditionals is useful in expert systems because it can be used to reason about imperfect generalizations, i.e., generalizations of the form: *In nearly all instances where α is true, β is true.* This paper defines a notion of importance and then shows that Adams' logic is equivalent to a logic for reasoning about generalizations of the form: *In all important instances where α is true, β is true.* A new method of proving results in Adams' logic — a method based upon the importance notion — is described. This method of proof is often quite easy, particularly since Venn diagrams can be used to help find proofs. An example using this new method of proof is presented. The example shows that Adams' logic justifies a common method of conflict resolution in expert systems, namely, the specificity-wins method. Finally, it is shown that Adams' logic can be used to reason about a topic seemingly unrelated to imperfect generalizations: the choices made by rational agents.

1 Introduction

1.1 Imperfect Generalizations

Expert systems and other rule-based decision aids reason with generalizations stored in their knowledge bases. Typically those generalizations are *imperfect* in that they have exceptions — exceptions that are rare. Thus, an imperfect generalization has the form: *In nearly all instances where α is true, β is true.*

Classical logic was constructed for reasoning with *perfect* generalizations — generalizations of the form: *In all instances where α is true, β is true.* Therefore, classical logic has limited utility for reasoning with imperfect generalizations. For example, classical logic provides no guidance about what to do when the conclusions implied by two imperfect generalizations contradict each other.

This has forced expert system developers to go beyond classical logic and to use their intuition for reasoning with imperfect generalizations. However, intuition is not always correct. It would be desirable then to have a formal logic for reasoning with imperfect generalizations.

1.2 Adams' Logic

Ernest Adams' logic of conditional statements [Adams, 1966, 1975, 1983, 1986] is such a logic. In Adams' logic, statements have the form $\alpha > \beta$ and are read "If α , then β ." Their meaning is that the conditional probability $Pr(\beta|\alpha)$ is close to one. It is possible, therefore, to interpret them as expressing the imperfect generalization: *In nearly all instances where α is true, β is true.* Thus, Adams' logic of conditionals can be used to reason about imperfect generalizations [Adams, 1974, Concluding Remark (1)].

For this reason, Adams' logic has been useful to researchers in artificial intelligence [Pearl, 1988, Section 10.2].

2 Entailment in Adams' Logic

Adams defined what it means, in his logic, for a collection of conditional premises to *probabilistically entail* either a single conditional conclusion [Adams, 1966, 1975] or a collection of alternative conditional conclusions [Adams, 1983].

I will not present here all the technical details needed to rigorously state Adams' definition of probabilistic entailment. For present purposes, the following loose characterization of probabilistic entailment will suffice. The conditional statement $\alpha > \beta$ will be said to be *reliable* whenever the conditional probability $Pr(\beta|\alpha)$ is close to one. Loosely speaking, the premises $\phi_1 > \psi_1, \dots, \phi_n > \psi_n$ probabilistically entail the alternative conclusions $\eta_1 > \mu_1, \dots, \eta_m > \mu_m$ if it is certain that at least one conclusion will be reliable whenever all the premises are reliable. (For a precisely stated version of the preceding, see Appendix A.)

If the premises $\phi_1 > \psi_1, \dots, \phi_n > \psi_n$ probabilistically entail the alternative conclusions $\eta_1 > \mu_1, \dots, \eta_m > \mu_m$, Adams' logic doesn't tell us which one of the conclusions will be reliable — only

Dist	Special
A-1	

that *at least one* of them will be reliable. Of course, if $m = 1$, then we know that the sole conclusion $\eta_1 > \mu_1$ must be reliable.

Algorithms have been developed for determining whether a specified collection of premises probabilistically entail a specified collection of alternative conclusions [Adams, 1983, 1986, Bamber, in press]. However, those algorithms are nonintuitive and provide little insight into the properties of probabilistic entailment.

The purpose of this paper is to present a new characterization of probabilistic entailment — one that provides insight.

3 Syntax

3.1 Statements

The syntax of Adams' logic will now be described. A conditional statement has the form $\alpha > \beta$ where α and β are statements from a propositional logic \mathcal{L} . Because statements in Adams' logic have this form, his logic will be denoted $\mathcal{L} > \mathcal{L}$.

Statements in \mathcal{L} are constructed as follows. Let S_1, \dots, S_k be primitive statements. Statements are constructed from primitive statements using the connectives: \neg (*not*), \wedge (*and*), \vee (*or*). Thus, all primitive statements are statements. If α and β are statements, then so are $(\neg\alpha)$, $(\alpha \wedge \beta)$, and $(\alpha \vee \beta)$. No other sequences of symbols are statements.

Let t (*true*) be an abbreviation for the tautology $(S_1 \vee (\neg S_1))$ and let f (*false*) be an abbreviation for the contradiction $(S_1 \wedge (\neg S_1))$.

3.2 Atoms

An atom is any statement T_k that is constructed as follows. Let T_1 denote *either* S_1 *or* $(\neg S_1)$. For $i = 2, \dots, k$, let T_i denote *either* $(T_{i-1} \wedge S_i)$ *or* $(T_{i-1} \wedge (\neg S_i))$. Thus, there are 2^k atoms in \mathcal{L} . Let \mathcal{A} denote the set of all atoms in \mathcal{L} . Let $\mathcal{P}(\mathcal{A})$ denote the power set of \mathcal{A} (i.e., the set of all subsets of \mathcal{A}).

If $\alpha, \beta \in \mathcal{L}$, let $\alpha \models \beta$ mean that the statement α entails (in propositional logic) the statement β .

For any statement $\alpha \in \mathcal{L}$, let $\mathcal{A}(\alpha)$ denote the set of atoms in \mathcal{L} that entail α . In other words,

$$\mathcal{A}(\alpha) = \{A \in \mathcal{A} : A \models \alpha\}.$$

Thus, $\mathcal{A}(f) = \emptyset$ and $\mathcal{A}(t) = \mathcal{A}$. If α and β are any statements in \mathcal{L} , then

$$\begin{aligned}\mathcal{A}(\neg\alpha) &= \mathcal{A} - \mathcal{A}(\alpha), \\ \mathcal{A}(\alpha \wedge \beta) &= \mathcal{A}(\alpha) \cap \mathcal{A}(\beta),\end{aligned}$$

and

$$\mathcal{A}(\alpha \vee \beta) = \mathcal{A}(\alpha) \cup \mathcal{A}(\beta).$$

Moreover, $\alpha \models \beta$ if and only if $\mathcal{A}(\alpha) \subset \mathcal{A}(\beta)$. Finally, if $\mathcal{A}(\alpha) = \{A_1, \dots, A_r\}$ where $r \geq 1$, then α is logically equivalent to $A_1 \vee \dots \vee A_r$.

4 Importance-wise Entailment

4.1 The Notion of Importance

Two closely related ideas will now be defined:

importance functions and important-part functions.

An importance function assigns an importance to every atom in \mathcal{A} . Specifically, an *importance function* is a function from \mathcal{A} into the integers. So, if $I(\cdot)$ is an importance function and if A is an atom, then $I(A)$ is either a positive integer, a negative integer, or zero. $I(A)$ is called the *importance* of A . The larger $I(A)$ is, the *more important* A is said to be. If $I(A) \geq 0$, then A is said to be *nonnegative*.

An important-part function divides every set of atoms into two mutually exclusive and exhaustive parts: an important part and an unimportant part. Specifically, an *important-part function* is a function $ip(\cdot)$ that maps $\mathcal{P}(\mathcal{A})$ into $\mathcal{P}(\mathcal{A})$ and has the following properties. For all $X, Y \subset \mathcal{A}$:

- (i) $ip(X) \subset X$.
- (ii) If $Y \subset X$ and Y intersects $ip(X)$, then $ip(Y) = Y \cap ip(X)$.
- (iii) If $Y \subset X$ and $ip(X) = \emptyset$, then $ip(Y) = \emptyset$.

$ip(X)$ is called the *important part* of X . The *unimportant part* of X is the part that remains after removing the important part.

Suppose that $I(\cdot)$ is an importance function and that $ip(\cdot)$ is an important-part function. Then $I(\cdot)$ is said to *generate* $ip(\cdot)$ if and only if, for every set of atoms $X \subset \mathcal{A}$,

$$\begin{aligned}ip(X) &= \{A \in X : (\forall B \in X) I(A) \geq I(B)\} \\ &\quad \cap \{A \in X : I(A) \geq 0\}.\end{aligned}$$

This equation says that $ip(X)$ consists of the most important nonnegative atoms in X .

It is straightforward to show that every importance function generates an important-part function and, conversely, every important-part function is generated by an importance function.

4.2 Definition of Importance-wise Entailment

A new type of entailment in $\mathcal{L} > \mathcal{L}$ will now be defined.

A statement $\alpha > \beta$ in $\mathcal{L} > \mathcal{L}$ is said to *describe* an important-part function $ip(\cdot)$ if and only if $ip[\mathcal{A}(\alpha)] \subset \mathcal{A}(\beta)$. Moreover, suppose that $I(\cdot)$ generates $ip(\cdot)$. Then, $\alpha > \beta$ is said to *describe* $I(\cdot)$ if and only if $\alpha > \beta$ describes $ip(\cdot)$. This means that $\alpha > \beta$ describes $I(\cdot)$ if and only if the most important nonnegative atoms in $\mathcal{A}(\alpha)$ are all contained in $\mathcal{A}(\beta)$.

Let \mathcal{P} and \mathcal{C} be finite collections of statements in $\mathcal{L} > \mathcal{L}$. The premises \mathcal{P} are said to *importance-wise entail* the alternative conclusions \mathcal{C} if and only if every *important-part function* that is described by all of the premises is also described by at least one conclusion. From this it follows that the premises \mathcal{P} importance-wise entail the alternative conclusions \mathcal{C} if and only if every *importance function* that is described by all of the premises is also described by at least one conclusion.

4.3 Equivalence with Probabilistic Entailment

Although importance-wise entailment and probabilistic entailment appear to be very different, they are equivalent. In other words, \mathcal{P} probabilistically entails \mathcal{C} if and only if \mathcal{P} importance-wise entails \mathcal{C} . [For a sketch of how this is proved, see Appendix B.]

Loosely speaking, the equivalence of probabilistic entailment and importance-wise entailment means the following. Statements of the form $\alpha > \beta$ can be interpreted in two different ways.

- (i) *In nearly all instances where α is true, β is true.*
- (ii) *In all important instances where α is true, β is true.*

Under either of these interpretations, $\alpha > \beta$ is a generalization having *limited* exceptions. Under interpretation (i), exceptions are limited in that they are *rare*. Under interpretation (ii), exceptions are limited in that they are *unimportant*. Under either interpretation, the inferences drawn in Adams' logic are correct.

5 Example of Importance-wise Entailment

5.1 Introduction

The equivalence between probabilistic entailment and importance-wise entailment can often be put to good use. It is often easy to demonstrate results in Adams' logic using proofs based upon important-part functions. This is justified because, whenever one wants to prove probabilistic entailment, it suffices to prove importance-wise entailment. Moreover, finding a proof of importance-wise entailment can often be made easier by drawing appropriate Venn diagrams. An example will now be given.

5.2 Conflict Resolution

The example shows that Adams' logic justifies a method of conflict resolution that is frequently used in expert systems. In fact, the example shows that the method can be generalized slightly.

Suppose that α , β , γ , and δ are statements in \mathcal{L} and that γ and δ are contradictory (i.e., $\gamma \models \neg\delta$). Suppose it is known that, when α is true, γ is nearly always true. Moreover, when β is true, δ is nearly always true. This knowledge is built into an expert system in the form of rules. One rule says that, in any instance where α is true, γ should be concluded. Another rule says that, in any instance where β is true, δ should be concluded. What should be done in an instance where both α and β are true? Both rules cannot be applied because, then, both γ and δ would be concluded — and γ and δ contradict each other. One common method of resolving this conflict is for the more specific rule to win [Rich, 1983, pp. 69-71, Winston, 1984, pp. 170-171]. For example, suppose that β is more specific than α . In other words, $\beta \models \alpha$. Thus,

in instances when β is true, α is *always* true. Because β is more specific than α , the rule involving β wins and, thus, δ is concluded rather than γ .

5.3 The Example

It will now be demonstrated, using importance-wise entailment, that Adams' logic justifies the specificity-wins method of conflict resolution. Moreover, Adams' logic justifies a slight generalization of this method of conflict resolution.

Recall that, when α is true, γ is nearly always true and that, when β is true, δ is nearly always true. These facts can be expressed in Adams' logic by the statements $\alpha > \gamma$ and $\beta > \delta$. Rather than assuming that when β is true α is *always* true, a weaker assumption will be made. It will be assumed that, when β is true, α is *nearly always* true. This can be expressed in Adams' logic by the statement $\beta > \alpha$.

Theorem. Suppose that α , β , γ , and δ are statements in \mathcal{L} and that $\gamma \models \neg\delta$. Then, in Adams' logic, the premises $\alpha > \gamma$, $\beta > \delta$, and $\beta > \alpha$ probabilistically entail the conclusion $(\alpha \wedge \beta) > \delta$.

Proof. Because probabilistic entailment and importance-wise entailment are equivalent, it suffices to prove importance-wise entailment.

Let $ip(\cdot)$ be any important-part function that is described by the premises. Because $ip(\cdot)$ is described by $\beta > \alpha$, $ip[\mathcal{A}(\beta)] \subset \mathcal{A}(\alpha)$. Similarly, because $ip(\cdot)$ is described by $\beta > \delta$, $ip[\mathcal{A}(\beta)] \subset \mathcal{A}(\delta)$. Finally, by Property (i) of important-part functions, $ip[\mathcal{A}(\beta)] \subset \mathcal{A}(\beta)$. The foregoing is illustrated in Figure 1.

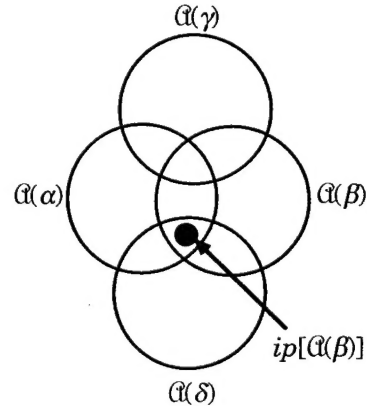


Figure 1. Sets of atoms.

It will now be shown that

$$ip[\mathcal{A}(\alpha) \cap \mathcal{A}(\beta)] = ip[\mathcal{A}(\beta)]. \quad (1)$$

If $ip[\mathcal{A}(\beta)] \neq \emptyset$, then $\mathcal{A}(\alpha) \cap \mathcal{A}(\beta)$ intersects $ip[\mathcal{A}(\beta)]$ and so, by Property (ii) of important-part functions, Equation (1) holds. On the other hand, if $ip[\mathcal{A}(\beta)] = \emptyset$, then by Property (iii) of important part functions, $ip[\mathcal{A}(\alpha) \cap \mathcal{A}(\beta)] = \emptyset$ and, thus, Equation (1) holds.

Recall that $\mathcal{A}(\alpha \wedge \beta) = \mathcal{A}(\alpha) \cap \mathcal{A}(\beta)$. Therefore,

$$ip[\mathcal{A}(\alpha \wedge \beta)] = ip[\mathcal{A}(\beta)] \subset \mathcal{A}(\delta).$$

Thus, $ip(\cdot)$ is described by the statement $(\alpha \wedge \beta) > \delta$. This demonstrates that $(\alpha \wedge \beta) > \delta$ is importance-wise entailed by the premises. ■

5.4 Monotonicity

Note that the proof of the above theorem makes no use of the premise $\alpha > \gamma$. The proof demonstrates that the two premises $\beta > \delta$ and $\beta > \alpha$ probabilistically entail the conclusion $(\alpha \wedge \beta) > \delta$. The reason this proof also demonstrates that the three premises $\alpha > \gamma$, $\beta > \delta$, and $\beta > \alpha$ probabilistically entail the conclusion $(\alpha \wedge \beta) > \delta$ is that Adams' logic has the property of *monotonicity*. What this means is that, if the collection of premises \mathcal{P} probabilistically entails the collection of alternative conclusions \mathcal{C} and if $\mathcal{P} \subset \mathcal{Q}$, then \mathcal{Q} probabilistically entails \mathcal{C} .

6 Rational Agents' Choices

6.1 Introduction

Surprisingly, not only can Adams' logic can be used to reason about imperfect generalizations, it can also be used to reason about a topic that seems totally unrelated to imperfect generalizations. Specifically, Adams' logic can be used to reason about the choices made by a rational agent.

By an *agent* is meant either a person, animal, or machine that makes choices. An agent is said to be *rational* if its choices are always consistent with its preferences.

6.2 Preferences

It is assumed here that an agent has stable preferences. For any two objects A and B , it is assumed that any agent would either prefer to be given A rather than B , or prefer to be given B rather than A , or be indifferent whether it was given A or B .

Suppose that \mathcal{O} is a set of 2^k objects. Let $o(\cdot)$ be a one-to-one function that maps \mathcal{A} onto \mathcal{O} . Because the function $o(\cdot)$ establishes a one-to-one correspondence between objects and atoms, it has the effect of labeling each object in \mathcal{O} with an atom in \mathcal{A} .

As a result of this one-to-one correspondence between objects and atoms, any agent's preferences among objects in \mathcal{O} can be represented — using a scheme to be described below — by an importance function $I(\cdot)$ on \mathcal{A} . Different agents will have different preferences among the objects in \mathcal{O} and, consequently, their preferences will be represented by different importance functions.

What does it mean for an agent's preferences among objects in \mathcal{O} to be represented by an importance function $I(\cdot)$? It means the following. If $A, B \in \mathcal{O}$, the agent prefers being given $o(A)$ rather than $o(B)$ if and only if $I(A) > I(B)$. The agent is indifferent between $o(A)$ and $o(B)$ if and only if

$I(A) = I(B)$. Moreover, the agent prefers being given nothing at all rather than being given $o(A)$ if and only if $I(A) < 0$ and, conversely, the agent prefers being given $o(A)$ rather than nothing at all if and only if $I(A) > 0$. Finally, the agent is indifferent between being given $o(A)$ and nothing at all if and only if $I(A) = 0$.

It is assumed that, for every importance function $I(\cdot)$, there is potentially an agent whose preferences can be represented by $I(\cdot)$. In other words, an agent's preferences could be anything.

6.3 When Choices are Rational

To say that an agent's choices are rational says nothing about the agent's preferences — those preferences could be anything. An agent's choices are rational when they are consistent with the agent's preferences — whatever those preferences may be.

Let $X \subset \mathcal{A}$ be any set of atoms — including possibly the empty set. Suppose that the agent is given the opportunity to either take *one* object from the set

$$o(X) = \{o(A) : A \in X\}$$

or to take nothing at all. Let the option of taking nothing be denoted by sq (for status quo). In other words, the agent is required to take *exactly one* member from the set $o(X) \cup \{sq\}$.

If the agent is rational, it will consult its preferences and will find the set of most preferred options in $o(X) \cup \{sq\}$. The agent will be indifferent between all options in that set. Hence, the agent will choose one option — at random — from the set of most preferred options. The agent's choice may vary from occasion to occasion, but it will always be a member of the set of most preferred options.

Suppose that an agent's preferences can be represented by the importance function $I(\cdot)$ and that $I(\cdot)$ generates the important-part function $ip(\cdot)$. Then, the set of most preferred options in $o(X) \cup \{sq\}$ will be either $ip(X)$ or $ip(X) \cup \{sq\}$.

6.4 Entailment and Rational Choice

A statement $\alpha > \beta$ in $\mathcal{L} > \mathcal{L}$ will be said to *describe* an agent's choices if and only if, whenever the agent is offered the choice of exactly one option from the set $o[\mathcal{A}(\alpha)] \cup \{sq\}$, the agent's choice always belongs to the set $o[\mathcal{A}(\beta)] \cup \{sq\}$.

A rational agent's choices will be described by $\alpha > \beta$ if and only if the agent's most preferred options in $o[\mathcal{A}(\alpha)] \cup \{sq\}$ are all members of $o[\mathcal{A}(\beta)] \cup \{sq\}$. But the set of most preferred options is either $o(ip[\mathcal{A}(\alpha)])$ or the union of the latter set with $\{sq\}$. Hence, $\alpha > \beta$ describes a rational agent's choices if and only if $ip[\mathcal{A}(\alpha)] \subset \mathcal{A}(\beta)$. In other words, $\alpha > \beta$ describes a rational agent's choices if and only if $\alpha > \beta$ describes the importance function $I(\cdot)$ that represents the agent's preferences.

From the above, it follows that Adams' logic can be used for reasoning about a rational agent's choices. Let \mathcal{P} and \mathcal{C} be finite collections of

statements in $\mathcal{L} > \mathcal{L}$. Suppose that a particular rational agent's choices are known to be described by every statement in \mathcal{P} . Then, it is a certainty that the agent's choices will be described by at least one statement in \mathcal{C} if and only if \mathcal{P} importance-wise entails \mathcal{C} . But importance-wise entailment is equivalent to probabilistic entailment.

6.5 A Potential Application

An as-yet-unexplored possibility for applying Adams' logic is the following. Any machine agent that interacts, either collaboratively or competitively, with other human or machine agents in a common environment needs to make inferences about the choices of actions made by the other agents. Although never designed for that purpose, Adams' logic might be used by machine agents for making such inferences. This might be useful, for example, in war-gaming simulators.

7 Summary

Unlike classical logic, Adams' logic can be used to reason about imperfect generalizations. This paper showed that entailment in Adams' logic can be understood in terms of the notion of importance. An example illustrating the use of important-part functions and Venn diagrams to demonstrate entailment showed that Adams' logic justifies a common method of conflict resolution in expert systems, namely, the specificity-wins method. Finally, it was shown that Adams' logic can be used to reason about the choices made by rational agents.

A Appendix: Probabilistic Entailment Theorem

The following theorem — for a proof, see [Bamber, in press, Theorem 7] — provides a precise statement of when a collection of premises probabilistically entails a collection of alternative conclusions. [In the following, it is understood that $Pr(\beta|\alpha) = 1$ whenever $Pr(\alpha) = 0$.]

Theorem. The premises $\phi_1 > \psi_1, \dots, \phi_n > \psi_n$ ($n \geq 0$) probabilistically entail the alternative conclusions $\eta_1 > \mu_1, \dots, \eta_m > \mu_m$ ($m \geq 0$) if and only if, for every $\epsilon > 0$, there exists a $\delta > 0$ such that, for every probability measure $Pr(\cdot)$ on \mathcal{L} , at least one of the conditional probabilities $Pr(\mu_1|\eta_1), \dots, Pr(\mu_m|\eta_m)$ will be $1 - \epsilon$ or larger provided that every one of the conditional probabilities $Pr(\psi_1|\phi_1), \dots, Pr(\psi_n|\phi_n)$ is $1 - \delta$ or larger.

B Appendix: Entailment Equivalence

This appendix sketches the proof that probabilistic entailment and importance-wise entailment are equivalent.

Definition. A p -ordering \leq is a weak ordering (i.e., a transitive connected relation) on \mathcal{L} that has

the following properties. For all $\alpha, \beta \in \mathcal{L}$:

- (i) If $\alpha \models \beta$, then $\alpha \leq \beta$.
- (ii) Either $(\alpha \vee \beta) \leq \alpha$ or $(\alpha \vee \beta) \leq \beta$.

A p -ordering is *degenerate* if $t \leq f$. Otherwise, it is *nondegenerate*.

Definition. Suppose $\alpha, \beta \in \mathcal{L}$. Then $\alpha > \beta$ describes the p -ordering \leq if and only if either $\alpha \leq f$ or $(\alpha \wedge (\neg\beta)) < (\alpha \wedge \beta)$.

Definition. A p -ordering \leq and an importance function $I(\cdot)$ are *related* if and only if they have the following properties. For all atoms $A, B \in \mathcal{A}$,

- (i) $A \leq f$ if and only if $I(A) < 0$.
- (ii) $A \leq B$ if and only if $I(A) < 0$ or $I(A) \leq I(B)$.

Adams [1986] proved the following theorem. It is easy to show that the theorem remains true if the word "nondegenerate" is removed from the statement of the theorem.

Theorem. Suppose that \mathcal{P} and \mathcal{C} are finite collections of statements in $\mathcal{L} > \mathcal{L}$. Then \mathcal{P} probabilistically entails \mathcal{C} if and only if every [nondegenerate] p -ordering that is described by every member of \mathcal{P} is also described by at least one member of \mathcal{C} .

The following two propositions are not hard to prove.

Proposition. Every p -ordering is related to some importance function and vice versa.

Proposition. Suppose that a p -ordering and an importance function are related. Let $\alpha, \beta \in \mathcal{L}$. Then $\alpha > \beta$ describes the p -ordering if and only if it describes the importance function.

From the theorem and two propositions above, it immediately follows that probabilistic entailment and importance-wise entailment are equivalent.

References

- [Adams, 1966] Ernest W. Adams. Probability and the logic of conditionals. Pages 265-316 in: Jaakko Hintikka and Patrick Suppes, editors. *Aspects of Inductive Logic*. North-Holland Publishing, Amsterdam, 1966.
- [Adams, 1974] Ernest W. Adams. The logic of 'almost all'. *Journal of Philosophical Logic*, 3: 3-17, 1974.
- [Adams, 1975] Ernest W. Adams. *The Logic of Conditionals*. D. Reidel Publishing, Dordrecht, Holland, 1975.
- [Adams, 1983] Ernest W. Adams. Probabilistic enthymemes. *Journal of Pragmatics*, 7: 283-295, 1983.
- [Adams, 1986] Ernest W. Adams. On the logic of high probability. *Journal of Philosophical Logic*, 15: 255-279, 1986.
- [Bamber, in press] Donald Bamber. Probabilistic entailment of conditionals by conditionals. *IEEE Transactions on Systems, Man, and Cybernetics*, in press.
- [Pearl, 1988] Judea Pearl. *Probabilistic Reasoning*

in Intelligent Systems. Morgan Kaufmann, San Mateo, California, 1988.

[Rich, 1983] Elaine Rich. *Artificial Intelligence.* McGraw-Hill, New York, 1983.

[Winston, 1984] Patrick Henry Winston. *Artificial Intelligence*, Second Edition. Addison-Wesley, Reading, Massachusetts, 1984.